DOCUMENT RESUME

ED 255 815 CG 018 147

AUTHOR Stevenson, John F.; Norcross, John C.
TITLE Local Training Evaluation: Who's Doing It and What Works?
PUB DATE Aug 84
NOTE 28p.; Paper presented at the Annual Convention of the American Psychological Association (92nd, Toronto, Ontario, Canada, August 24-28, 1984). Best copy available.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Clinical Experience; *Clinical Psychology; Course Evaluation; *Evaluation Methods; Evaluation Needs; *Formative Evaluation; *Professional Education; *Program Effectiveness; *Self Evaluation (Groups); Student Evaluation

ABSTRACT

Local evaluations of clinical training programs in psychology are generally conducted from a formative perspective, necessitating careful examination of the training structure, process, and results. To learn about and document the extent of local training evaluation enterprises, two national surveys that solicited general program information, information about the nature and frequency of use of various methods for evaluating training, subjective assessments of evaluation impact, and obstacles to evaluation were mailed to directors of psychology training clinics and to university directors of clinical psychology. The responses (N=87; N=67) indicated that the most frequently used and highest rated source of qualitative evidence of training impact on students was the supervisors' oral evaluations of clinical trainees. The most popular, systematic, qualitative sources of evidence of student performance included supervisors' written evaluations (84 percent) and internship supervisors' reports (84 percent). Some quantitative sources of evidence of student performance used by a minority of respondents included rate of student acceptance to first-choice internships (43 percent), knowledge tests (40 percent), and supervisors' quantitative ratings of students (40 percent). Qualitative means of assessing programs were supervisors' written evaluations and American Psychological Association accreditation reports. Quantitative evidence of program effectiveness included students' course evaluations and ratings of clinical supervisors. The pattern of results attested to the extensive, highly variable nature of current training evaluation activities. Inadequate resources (time, money, personnel) to conduct meaningful evaluation were perceived as the most serious source of difficulty for both clinical and program directors. (NRB)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

BEST COPY AVAILABLE

ED255815

# Local Training Evaluation:

# Who's Doing It and What Works?

John F. Stevenson and John C. Norcross

University of Rhode Island

Presented at the American Psychological Association Annual Convention,

Toronto, August, 1984.

CG 018147

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John F Stevenson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

## Prologue

After spending a year on leave conducting evaluation research in a psychiatric hospital, I returned to my department just as a new clinical training grant was being implemented. The idea of evaluating the training and treatment provided under the auspices of the grant seemed a natural extension of my interests. I looked forward to designing and conducting a sophisticated and useful evaluation program, based in our training clinic. Six years later, I can safely say I was naive. Although I am much more skeptical now about such an enterprise, I am still intrigued by the possibilities and still convinced of the merits of the undertaking. Disturbed by the difficulties we encountered locally, my coauthor and I have sought out evaluation experts in other clinical programs for collaboration and support. To our chagrin, we have found little expertise, some support, quite a bit of cynicism, and a number of curious folks hoping we could tell them something.

Today we will discuss what we have learned about local evaluation of clinical training -- what is being done, what obstacles have been encountered, what approaches seem promising, and what the future may hold. We will draw on two national surveys as well as our own experiences.

## Significance

Later in this symposium, much will be said about the national importance of the evaluation of clinical training. A national strategy, to be described by Lee Sechrest, is evolving in response to the significant need to inform and influence policy. Why bother at all with evaluation of local training sites? Why not restrict evaluation to the national level, where adequate resources, large samples, technical expertise, and an integrative, policy-oriented perspective seem more available?

Local training evaluation has several important justifications. First, methods and measures developed at the local level may be aggregated to address national concerns. Second, many practical issues for local trainers are too site-specific to benefit from national studies. Third, locally generated feedback on training effectiveness is more likely to influence local decisions. Fourth, local programs are becoming increasingly concerned about accountability and funding, and evaluation data can be used to attract and justify economic support from a widening range of sources. Fifth, students should be exposed to responsible professional practices, and training evaluation provides a model for the evaluation of services. Sixth, local programs can serve as greenhouses for the development and dissemination of innovative training methods, and this process requires careful evaluation.

## The Self-Evaluating Organization

An important distinction is made in the literature between "formative" and "summative" evaluation. Formative evaluation is conducted as a rational means of guiding program development and improvement. Summative evaluation, usually performed by an external evaluator, is conducted to determine the overall merit of a program and may lead to program termination. Differing evaluation questions and strategies are associated with these two orientations. Local evaluation is generally conducted from a formative perspective, necessitating careful examination of the training structure, process, and results. Unlike the classic experiment, with predefined treatment and control groups and, usually, a limited time perspective, formative evaluation is an ongoing, longitudinal process embedded in the organization. Experiments may be designed, for example, to compare distinct training methods, but quasi-experimental and qualitative approaches predominate.

One way of characterizing a formative perspective, useful for considering current and future training evaluation activities, is in a general systems framework. Figure 1 displays an outline of such a framework and Figure 2 suggests relevant sources of evidence. Note that in addition to a variety of types of evidence regarding the impact of training on individual students, information on the content, structure, pattern, and timing of student exposure to training experiences is included. A number of economic measures are also possible. Within a systems framework formative evaluation is represented as one or more feedback loops, providing information to decision makers on the functioning of training components as well as on the ultimate impact of training on students. Another way of saying this is that evaluation is an indispensible part of an organization's "reality testing".

## National Survey Findings: Local Evaluation Activity

Our own efforts to improve the reality testing of the clinical program and psychology clinic at the University of Rhode Island are described in a recent article (Stevenson, Norcross, King & Tobin, 1984). After some modest success and considerable frustration, we decided to investigate the experiences of fellow local evaluators, hoping to learn from more mature evaluation enterprises and to document their extent.

In the past two and a half years, John Norcross and I have completed two national surveys of local training evalution activity. One of these was mailed to the 131 persons on the mailing list of the Association of Directors of Psychology Training Clinics (ADPTC). The return rate for this survey was 66% ($\underline{N}$=87). The second survey was mailed to the 109 American members of the Council of University Directors of Clinical Psychology. The return rate was 61% ($\underline{N}$=67), representing roughly one half of the 124 doctoral clinical

psychology programs accredited by APA.

Both survey questionnaires solicited general program information, nature and frequency of use of various methods for evaluating training, subjective assessment of evaluation impact, and obstacles to evaluation. A few additional questions were unique to each survey. Detailed descriptions of the surveys and findings are provided elsewhere (Norcross & Stevenson, 1984; Stevenson & Norcross, in press) and can be obtained upon request from either author.

Tables 1-5 summarize survey findings regarding the frequency of routine use of, and amount of importance attached to, a number of sources of information about training program effectiveness. Most of the data come from our survey of directors of clinical training (the "Program" columns). Where comparable data were obtained in our survey of directors of training clinics, these are also provided (the "Clinic" column). The tables group sources of information on the basis of three dimensions: (1) focus of evaluation (aggregating performance indicators for individual students versus assessing the training program directly); (2) type of evaluation data (qualitative or quantitative); and (3) method of data collection (impressionistic or systematic).

Now to our first question about local evaluation of clinical training: "Who's doing it?" Since everyone is evaluating training in some fashion, a more refined version of this question might be "Who's doing what?" Looking at the frequencies of use for various categories of evaluation information reveals what's going on around the country.

Qualitative, impressionistic evidence of training impact on students (Table 1) is widely used and judged relatively important in forming an overall impression of program quality. Supervisors' oral evaluations of clinical

trainees are the most frequently used (76%) and highest rated (4.1) in this group

Qualitative, systematic sources of evidence of student performance (Table 2) vary greatly in frequency of use, but most are rated as relatively influential by those who do use them. Most popular in this category are supervisors' written evaluations (84%) and internship supervisors' reports (84%). These two qualitative sources are used more frequently than any other source included in the survey, and supervisors' written evaluations are rated the most important source of overall impressions of training quality. Clearly, verbal evidence bearing directly on clinical performance of students plays a central role in directors' judgments of their training programs. It is not so clear how such qualitative indicators are aggregated and monitored to yield decisions about program improvements.

Quantitative evidence of student performance might lend itself better to longitudinal monitoring of program effectiveness and quasi-experimental comparisons of various stages and components of training. As seen in Table 3, some quantitative sources are used by a substantial minority of respondents, but none by a majority. The rate of student acceptance to first-choice internships (43%), knowledge tests (40%), and supervisors' quantitative ratings of students (40%) are most used and most influential in this category. These indices appear to combine the advantages of quantitative precision with relevance to training objectives.

Program-focused means of assessing effectiveness move away from evaluation of individual student performance to sources of evidence about the program itself. This is particularly important in formative evaluation because of the need to link outcomes (e.g., student skill increments) with elements of the training program (e.g. separate therapy training components, stages in

training, alternative supervision modalities, apprenticeship versus collaborative research training models). Furthermore, some aspects of training are so evidently important that direct monitoring seems essential -- for example, supervision hours and quality, amount of exposure to particular modalities, enrollment patterns in required courses.

Qualitative means of assessing programs (Table 4) are intermediate in frequency of usage and rated importance. APA accreditation reports perhaps deserve a category all their own, since they combine many forms of evaluative input into a qualitative summary, and since they are produced by outsiders and serve both accountability and feedback functions. These reports are frequently used (77%) and rated as very influential (4.3) in forming impressions of overall program quality. They are second only to supervisors' written evaluations in rated importance. Other influential qualitative sources include curriculum reviews with student input and site visits to off-campus training sites.

Quantitative evidence of program effectiveness (Table 5) includes some very popular sources, such as students' course evaluations (70%) and ratings of clinical supervisors (68%). These are seen as relatively influential in judging program performance. Other quantitative sources are infrequently used and not very influential. Consumer feedback, in the form of student ratings of program effectiveness during and following training, is not common, and use of faculty ratings is rare. Our own experience suggests that a good deal can be learned when students and faculty are given the opportunity to provide anonymous input (both quantitative and qualitative) on various aspects of the training program. We have found substantial agreement between these two sources on both strengths and weaknesses. However, we have also found that the local politics of evaluation are brought to the fore when training and trainers, rather than students, are directly evaluated.

For the most part the responses of clinic directors serve to validate the more extensive information provided by program directors. When discrepancies between the two sources appear, they are generally understandable in the light of the differences in scope and focus of the two contexts. Clinical programs, for instance, are much more likely than training clinics to use tests of factual knowledge (40% versus 16%), probably because clinics are not primarily focused on imparting factual information to students.

We were also interested in the extent to which the various means of evaluating training had been employed in formal evaluation research designs. Table 6 summarizes the responses of program directors when asked whether they had ever done such a study. Although a number of programs have conducted more formal quantitative research, it is not perceived as influential. Least popular, but most influential, are longitudinal studies in which students are pre-tested and then re-tested following training.

The pattern of results attests to the extensive, highly variable nature of current training evaluation activities. Qualitative, student-focused sources are popular and influential, as might be expected, but many programs also use various kinds of quantitative evidence. A mix of qualitative and quantitative approaches is probably most useful. Innovative and promising means of measuring student competence are being tried in several places. Some program-focused sources are widely used, but much more can be done to explore how our training programs affect students, and where improvements are necessary. Relatively few programs are engaged in carefully designed quantitative research, and we now turn to a number of obstacles to formal evaluation which may help to account for this.

## Obstacles to Meaningful Evaluation

Our two surveys included questions regarding barriers to meaningful evaluation research. Training clinic directors were asked to list up to three major obstacles. One the basis of their replies, 12 categories were created for tabulating frequencies of spontaneously generated obstacles. These categories were then employed in our survey of directors of clinical training, who were asked to rate the extent of difficulty caused by each (1 = not an obstacle, 5 = very serious obstacle).

Table 7 presents the responses from both surveys, with the 12 categories ordered from the most serious to the least serious on the basis of mean ratings by program directors. Time restraints and inadequate methods and measures are seen by both groups as relatively serious obstacles to meaningful evaluation. Program directors are more concerned about lack of personnel and practical difficulties associated with collecting data from transient students than are clinic directors. The latter are more concerned about general apathy and disinterest. Both sets of respondents viewed inadequate funding for evaluation as a moderately serious obstacle.

To provide a more comprehensive picture, obstacles were grouped into four major classes, and mean ratings by program directors, as well as total endorsement frequencies by clinic directors, were calculated. Inadequate resources (time, money, personnel) to conduct meaningful evaluation are perceived by both groups as the most serious source of difficulty. Organizational resistance is second in importance for clinic directors, while pragmatic and technical problems are second for program directors. Ethical difficulties are not seen as much of a hindrance by either group.

These problems are familiar to anyone who has attempted to develop a training evaluation program. They are certainly a part of our experience at

the University of Rhode Island. The groupings we have provided help to represent the range of issues, but our experiences suggest that the obstacles are interconnected. One example of this is the probable connection between resource scarcity and organizational resistance. Resources are allocated according to priorities, and priorities are set to meet pressing needs and to avoid discomfort. Evaluation research not only requires resources, but it also shape priorities for their reallocation -- for this reason it may be resisted as a destabilizing process. A second example of the probable interdependence of obstacles is the connection between theoretical disparities among faculty, on the one hand, and inadequate measurement technology, on the other. Defining measurable objectives for training is made doubly difficult when measures must fit the myriad of professional and clinical agendas of faculty as well as traditional psychometric requirements. The principle of academic freedom must be balanced against the need to establish coherent program goals, and this sensitive situation is made painfully obvious when one attempts to empirically define objectives. In our survey of directors of clinical training, very few provided clear, operationalizable training goals, and this is problematic at our own institution. Underlying local differences of opinion about training goals are broader issues in the discipline concerning the nature of the profession and the ingredients of competent clinical performance.

Another methodological obstacle associated with broader concerns of the discipline is the difficulty of specifying the training intervention. Students are exposed to a variety of experiences in graduate clinical training, and it may be difficult to disentangle their effects. Furthermore, components, requirements, and personnel in the training program may change

frequently over time. However, the local effort to identify and clarify the significant components of the training process can have enormous utility for national efforts to determine the experiences provided in the doctoral training of clinical psychologists.

To restate the primary point, several local obstacles reflect discipline-wide issues, and local attempts to overcome these obstacles may make a significant contribution far beyond improvements in an individual program.

Impediments to evaluation are not only present before evaluation takes place, when resource allocation and planning decisions are made. In our own program, resistance during ongoing evaluation in the form of missing data is a serious problem. Because the numbers are small, any missing data lead immediately to concern about sample bias. We have found that return rates generally suffer when little public use is made of the data. This leads to acknowledgement of a final important obstacle: non-utilization of findings. Even when data are collected, results reported, and recommendations made, there may be little impact on policy. Our own experience is consistent with the well documented view that policy impact is most probable when policy makers genuinely want information in the first place and take an active, personal role in devising evaluation questions and reviewing results (Stevenson & Ciarlo, 1982). There isn't much point in asking questions to which no one wants the answer. In our view, recognition of evaluation as a formative feedback process, as described earlier, can improve utilization, even if it detracts from the prestige associated with antiquated formulations.

## What Works: Promising Approaches to Meaningful Evaluation

We will draw on the preceding sections to generate some suggestions for

those who keep tilting at this windmill. We are skeptical about universally applicable high-technology solutions. Advances in measurement and research methods must be coupled with local determination, inventiveness, and flexibility. Economic resource limitations need not constitute impervious obstacles if local leadership respects and supports evaluation in other ways. Our naivete has grudgingly given way to realism: an appreciation for the contingencies, difficulties, and resistances inherent in the task of evaluating clinical training. The commingling of nomothetic empiricism with the idiographic, qualitative nature of clinical training may be viewed as meddling or as instructing - as a threat to the power structure or as one means toward enlightened exercise of power, as creative applied science or as glorified bookkeeping. The view one takes of the evaluation enterprise is likely to influence its outcome.

Our formative, pragmatic approach to solving evaluation problems emphasizes the harnessing of technical means to meet practical needs. Rigor is important, not as an end in itself, but as a way of assuring trustworthy answers to the real questions. Cleverly devised measures of clinical skills are no substitute for faculty consensus on what the important skills are. Knowing what actually happens to students during training is as essential as knowing that students perform differently at the end of training. The following are some suggestions based on this perspective.

Making evaluation useful

The director of training and the evaluator are frequently the same person. Even when they are, each may not talk to the other enough! It is important for evaluation to address questions in such a way that the results can usefully inform decision and action. Judging by the survey responses,

program directors rely heavily on qualitative means of assessing their programs. Acquiring qualitative information more systematically might provide a useful complement to quantitative evidence. For example, semi-structured interviews with a random sample of each cohort of students might offer a clearer picture than casual hall conversations or anonymous rating scales. Some kinds of quantitative information are apparently valued when available--these include practicum and internship supervisors' ratings and tests of clinical knowledge. Integration of qualitative and quantitative feedback enhances the value of both.

To make information relevant for decision-making, it is important to link training objectives to specifiable elements of the training process. If students are expected to be skilled at conducting intake interviews, client assessments, case formulation, and termination, as well as specific intervention strategies, then feedback of supervisors' evaluation of these skills will be more influential if each skill is tied to specified training experiences. If master's theses are to be defended by the end of the third year, data indicating largescale failure to meet this goal will be more useful if the means for students to achieve the goal have been specified. In order to improve our understanding of the URI training program, we have begun conducting surveys of prospective, current, and former students. We are interested in what students' training experiences have been; how they evaluate them; what their goals and career directions are; what strengths, weaknesses, and changes they pinpoint; and how these things relate to such factors as gender, year of entry, and Boulder vs. Vail orientation. We hope survey findings will clarify the ties between the training process and outcomes. To examine our admission process we have recently conducted a survey of the (very

few!) students who were admitted to our program but chose to go elsewhere. Since our screening interview is a key means of informing prospective students about the program as well as aiding in admission decisions we were particularly concerned about the impact on students of this initial contact. From a systems perspective (Fig. 1) pre-admission exchanges begin the flow into and through the system.

We have also undertaken anonymous surveys of our faculty, with promising results. In academic institutions, the need to foster a close tie between evaluation and decision making calls for faculty involvement, since each faculty member exercises a fair amount of autonomous control over training. Definitions and measures of training objectives will have little value if they are perceived as irrelevant by the faculty who provide the training and ultimately determine the direction of the program. In our experience faculty support is essential for data collection as well as utilization of results. Therefore, it is important to generate faculty input on training objectives, curriculum issues, and needed changes.

As the discussion of a formative, systems-oriented framework suggested, we believe evaluation should be an integral, ongoing part of normal program operation rather than a one-time study. Our survey of clinical directors provides empirical support for this view in that ongoing evaluation of training was significantly more likely to influence reported policy impact. Our program has an Evaluation Team with practicum course credit available, and a number of the programs we surveyed report thesis research associated with self-evaluation activities. Another resource-saving tactic for integrating evaluation into normal operation is to make use of existing information channels. Supervisors may already provide written evaluations of students --

it is relatively easy to add some structure, and perhaps a few rating scales, to this evaluation. We are now in the process of extending our feedback system to include quantitative ratings of students on a variety of skill dimensions by internship supervisors, another readily available source.

Measuring training outcomes

The technology for measuring training outcomes is rapidly expanding. Scoring systems for videotaped, audiotaped and written work samples, simulation exercises, and client rating forms are all under development or available. Simulated licensure exams are in use, as are quantitative rating forms for supervisors. Validating these measures is a challenge. Since the measurement of therapy outcome itself is not yet well standardized, assessment of therapeutic skill is bound to be controversial. One implication of this is that multiple measures and multiple sources should be used to capture the range of skill dimensions. Another implication is that post-training internship and job performance measures can add an important perspective. A third implication is that local definitions of skill objectives have a legitimate role in selecting and validating measures for local use.

A pet interest of mine is to combine research on clinical decision making with the development of training evaluation methodology. It seems to me that a clearer understanding of the cognitive problem-solving processes of skilled and unskilled clinicians can simultaneously improve our theoretical understanding of clinical skills and increase our ability to measure and improve the impact of training. In psychiatry a beginning has already been made on this line of research (Hurcombe & Gallagher, 1984).

Beyond the local level

One reason for skepticism about the future of training evaluation is our

conviction that commitment beyond the local level is necessary. Local priorities and attitudes are affected by national priorities and attitudes. Local determination and inventiveness can be bolstered by the experiences of evaluators in other settings. One hope we have for this symposium is that it will lead to more active collaboration across settings, adding robustness and generalizability to indigenous, low-cost evaluation programs. Sharing of methods and measures, as well as collaborative research designs, seem likely to benefit all of us.

Another hope we have for this symposium is that a clearer national mandate for training evaluation will emerge, with a recognition that local self-evaluation can be encouraged through the promulgation of concrete standards and the dissemination of current methods. A clearinghouse for locally developed ideas would help. We are not talking about accountability requirements, which are notorious for their deadening effect on genuine evaluation, but about visible support for the old-fashioned belief that clinical psychology is a discipline whose training sites can provide a context in which research, training, and practice can enrich one another. The training of clinical psychologists is a costly business, and the health needs of the country present a demanding challenge. In such a context, careful evaluation of local training processes and outcomes can make a vital contribution.

## References

Norcross, J.C., & Stevenson, J.F. (1984). How shall we judge ourselves: Training evaluation in clinical psychology programs, Professional Psychology: Research and Practice, 15, 497-508.

Nurcombe, B., & Gallagher, R.M. III (1984). Clinical reasoning in psychiatry: Research and education. In B. Nurcombe and R.M. Gallagher, III (Eds). The clinical process in psychiatry. New York: Cambridge University Press.

Stevenson, J., & Ciarlo, J. (1982). Enhancing the utilization of mental health evaluation at state and local levels. In J. Stahler and W. Tash (Eds.), Innovative approaches to mental health evaluation. New York: Academic Press.

Stevenson, J.F., Norcross, J., King, J.T., & Tobin, K.G. (1984). Evaluating clinical training programs: A formative effort. Professional Psychology: Research and Practice, 15, 218-229.

Stevenson, J.F., & Norcross, J.C. (in press). Evaluation activity in psychology training clinics: National survey findings. Professional Psychology: Research and Practice.

Figure 1.

A Systems Framework for Clinical Training Program Evaluation: Content

| INPUT | THROUGHPUT | OUTPUT |
|---|---|---|
| Student characteristics at admission (knowledge, skills, values, demog) | Decision-making process | Student characteristics prior to internship and at graduation |
| | Admissions process | |
| | Curriculum content and structure | Post-graduation performance of students |
| | Enrollment patterns | Student satisfaction |
| | Time to completion of requirements toward degree | Program reputation |
| | Clinical skills training (assessment, supervised therapy, etc.) | |
| | Research training | |
| | Mechanisms for evaluating student competence and progress | |
| | Mentoring process | |
| | Financial support allocation | |

Figure 2.

A Systems Framework for Clinical Training Program Evaluation: Assessment

| INPUT | THROUGHPUT | OUTPUT |
|---|---|---|
| Archival admissions data | Archival data on courses, enrollments, student progress | Student questionnaires |
| Student questionnaires | Course evaluations | Specialized tests of clinical competency |
| Specialized clinical competency pretests | Records of supervision, therapy hours, modalities | Internship supervisor feedback |
| | Students' evaluations of supervisors | Internship and job placements |
| | Site visits | Licensure |
| | Supervisors' evaluations of students | Publications |
| | Student and applicant questionnaires | Applicant surveys |
| | Publications and presentations | |
| | Faculty assessment of training | |

Table 1

Student-focused Means of Assessing Clinical Training Effectiveness:

Qualitative, Impressionistic Sources

| Source of Evidence | Programs | | Clinics |
|---|---|---|---|
| | Frequency | Importance | |
| Supervisors' oral evaluations | 76% | 4.1 | 91% |
| Impressions of student growth | 73% | 3.9 | -- |
| Quality of internships students obtain | 71% | 4.0 | -- |
| Quality of 1st professional positions | 48% | 3.8 | -- |
| Reputations of former students among professional peers | 35% | 3.9 | -- |

Note. Importance in forming an overall impression of the quality of the program was rated from 1 ("not important") to 5 ("very important").

Table 2

Student-focused Means of Assessing Clinical Training Effectiveness: Qualitative, Systematic Sources

| Sources of Evidence | Programs | | Clinics |
|---|---|---|---|
| | Frequency | Importance | |
| Aggregated supervisors' written evaluations | 84% | 4.5 | 82% |
| Internship supervisors' reports | 84% | 4.2 | -- |
| Clinical case reports | 52% | 4.0 | -- |
| Case presentations | 35% | 3.8 | -- |
| Videotaped therapy sessions | 20% | 3.9 | 31% |
| Simulated cases | 11% | 3.3 | 18% |

Note. Importance in forming an overall impression of the quality of the program was rated from 1 ("not important") to 5 ("very important").

Table 3

Sudent-focused Means of Assessing Clinical Training Effectiveness: Quantitative Sources

| Sources of Evidence | Programs Frequency | Programs Importance | Clinics |
|---|---|---|---|
| Rate of student acceptance by 1st choice internship | 43% | 3.5 | -- |
| Knowledge tests | 40% | 4.2 | 16% |
| Supervisors' ratings of students | 40% | 4.1 | -- |
| Rate of licensure | 19% | 3.2 | -- |
| Simulated licensure exam | 8% | 3.3 | -- |
| Students' self-ratings | 3% | 3.0 | -- |

Note. Importance in forming an overall impression of the quality of the program was rated from 1 ("not important") to 5 ("very important").

Table 4

Program-Focused Means of Assessing Clinical Training Effectiveness: Qualitative Sources

| Sources of Evidence | Programs | | Clinics |
|---|---|---|---|
| | Frequency | Importance | |
| APA accreditation reports | 77% | 4.3 | -- |
| Curriculum reviews with student input | 56% | 4.1 | -- |
| Impresions of student satisfaction | 53% | 3.6 | -- |
| Site visits to off-campus training sites | 40% | 4.1 | -- |
| Reputation of your program | 35% | 3.5 | -- |
| Type and number of course offerings | 34% | 3.6 | -- |

Note. Importance in forming an overall impression of the quality of the program was rated from 1 ("not important") to 5 ("very important").

Table 5

Program-Focused Means of Assessing Clinical Training Effectiveness:

Quantitative Sources

| Sources of Evidence | Programs | | Clinics |
|---|---|---|---|
| | Frequency | Importance | |
| Students' ratings of courses | 76% | 4.1 | -- |
| Students' ratings of supervisors | 68% | 4.2 | 67% |
| Length of time to complete degree | 39% | 3.3 | -- |
| Students' ratings of satisfaction | 26% | 3.7 | -- |
| Students' ratings of training program effectiveness (during training) | 16% | 3.7 | -- |
| Students' ratings of training program effectivess (after completion) | 8% | 3.4 | -- |
| Faculty ratings of program effectiveness | 6% | 3.0 | -- |

Note. Importance in forming an overall impression of the quality of the program was rated from 1 ("not important") to 5 ("very important").

Table 6

Formal Designs for Training Evaluation

| Design | Frequency | Importance |
|---|---|---|
| Longitudinal (e.g., pre-post training) within group comparison | 26% | 3.2 |
| Cross-sectional comparison of new students with those who have already receiving training | 39% | 2.8 |
| Comparison to students in other programs | 29% | 2.8 |

Note. Importance in forming an overall impression of the quality of the program was rated from 1 ("not important") to 5 ("very important").

Table 7

Obstacles to Meaningful Training Evaluation

| Obstacles | Program Ratings (M) | Clinic Frequencies (n) |
|---|---|---|
| Time restraints | 3.2 | 26 |
| Inadequate methods and measures | 3.0 | 19 |
| Lack of personnel | 2.9 | 9 |
| Transience of students | 2.8 | 6 |
| Inadequate funding | 2.6 | 15 |
| General apathy and disinterest | 2.4 | 20 |
| Faculty resistance | 2.2 | 12 |
| Lack of belief in value of research evidence | 2.1 | 3 |
| Heterogeneity of students | 2.0 | 5 |
| Student resistance | 1.9 | 5 |
| Theoretical disparities among faculty | 1.9 | 12 |
| Ethical considerations | 1.8 | 2 |

Note. Mean program ratings are based on a 1-5 scale (1=not an obstacle; 5=very serious obstacle). Clinic frequencies are rates of spontaneous generation, when directors were asked to list up to three major obstacles.

Table 8

Major Obstacle Clusters

| Obstacles | Program Ratings (M) | Clinic Frequencies (n) |
|---|---|---|
| Resources (time, money, personnel) | 2.9 | 52 |
| Organizational resistance (apathy, resistance, theoretical disparities) | 2.2 | 47 |
| Pragmatic and technical problems (subject resistance and sampling difficulties, inadequate methods and measures) | 2.4 | 39 |
| Etnical considerations | 1.8 | 2 |

Note. Mean program ratings are based on a 1-5 scale (1=not an obstacle; 5=very serious obstacle). Clinic frequencies are rates of spontaneous generation, when directors were asked to list up to three major obstacles.